NASA TECHNICAL NOTE

NASA TN D-6288

C. 1

NASA TN D-6288

ACCEPTANCE SAMPLING FOR THE NORMAL DISTRIBUTION: STANDARD DEVIATION UNKNOWN BUT ASSUMED CONSTANT

by Steven M. Sidik

Lewis Research Center

Cleveland, Ohio 44135

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION • WASHINGTON, D. C. • MARCH 1971

1. Report No. NASA TN D-6288	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle ACCEPTANCE SAMPLING FOR THE NORMAL DISTRIBUTION: STANDARD DEVIATION UNKNOWN BUT ASSUMED CONSTANT		5. Report Date March 1971
		6. Performing Organization Code
7. Author(s) Steven M. Sidik		8. Performing Organization Report No. E-6034
		10. Work Unit No. 129-04
9. Performing Organization Name and Address Lewis Research Center National Aeronautics and Space Administration Cleveland, Ohio 44135		11. Contract or Grant No.
		13. Type of Report and Period Covered Technical Note
12. Sponsoring Agency Name and Address National Aeronautics and Space Administration Washington, D.C. 20546		14. Sponsoring Agency Code
15. Supplementary Notes		

16. Abstract

Often an estimate of the standard deviation based on existing data is used as a fixed value in a series of standard deviation known sampling plans. It is shown that this causes the operating characteristic curve of the standard deviation known plan to become a random variable. Its distribution depends upon the distribution of the estimate of the standard deviation. A method is then presented for computing confidence limits on the attained curve. It is also shown how to solve the converse problem of attaining specified limits on the curve for any given true proportion defective.

17. Key Words (Suggested by Author(s)) Acceptance sampling; Quality control; Industrial quality control		18. Distribution Statement Unclassified - unlimited	
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 21	22. Price* $3.00

*For sale by the National Technical Information Service, Springfield, Virginia 22151

ACCEPTANCE SAMPLING FOR THE NORMAL DISTRIBUTION: STANDARD DEVIATION UNKNOWN BUT ASSUMED CONSTANT

by Steven M. Sidik

Lewis Research Center

SUMMARY

Standard deviation known plans can achieve a considerable reduction over standard deviation unknown plans in the number of sampled items required. One procedure often used is that of computing an estimate of the standard deviation based upon previously accumulated data. This estimate is then used as the known value in a standard deviation known plan. The estimate is a random variable, however. It is shown that this causes the operating characteristic curve to be a random variable also and its distribution depends on the distribution of the estimate of the standard deviation.

Since the operating characteristic curve is a random variable, it is of interest to consider the two problems of computing confidence limits on the curve given a random sample of data and of determining a sample size to achieve specified confidence limits. Solutions for each of these problems are presented. A numerical example is also used to illustrate the solution to the second problem.

INTRODUCTION

Variables acceptance sampling plans have been widely used in industry and government. This statistical quality control technique is used to maintain a minimum standard of quality for producers or consumers of goods.

Assume that some item is produced, shipped, or received in lots of size M. Also suppose that associated with each item there is a critical variable which may be used as a quality characteristic. This may be some such thing as length, width, voltage, and so forth.

This critical variable is most commonly assumed to be normally distributed. That is, $x \sim N(\mu, \sigma^2)$. This report studies the situation where μ is assumed unknown and constant within a lot but permitted to vary from lot to lot and σ^2 is assumed unknown but is constant over all lots.

Any item which has the critical variable $X > X_U$ is called a defective that exceeds the upper specification limit X_U. Likewise, if $X < X_L$, the item is said to be a defective falling below the lower specification limit X_L. A variable may be restricted to lie between both an upper and a lower limit in which case we say it has double specification limits.

Consider figure 1 where density functions for two normal distributions with the same variance but differing means are illustrated. The top graph of figure 1 shows $x \sim N\left(\mu_{p_2}, \sigma^2\right)$. There is a proportion p_2 of the items in a lot following this distribution which are defective. This proportion is called the consumers quality point and it is assumed the consumer would like the probability of accepting lots with this proportion defective or more to be less than or equal to β. The bottom graph of figure 1 shows $x \sim N\left(\mu_{p_1}, \sigma^2\right)$. The proportion p_1 is referred to as the producers quality point and it is assumed the producer would like the probability of accepting lots with this proportion defective or less to be greater than or equal to $1 - \alpha$. Most generally, α is about 0.01 or 0.05 and β is about 0.10 or 0.05. The values p_1 and p_2 are mutually agreed upon by consumer and producer in some manner.

Given a lot of M items, we wish to sample n such items randomly and decide whether or not too large a proportion of the lot falls outside the permissible range. This range may be limited by the upper limit X_U, lower limit X_L, or both an upper and lower limit. The sample size and decision procedure should be such that the producer and consumer are satisfied. That is, lots with a proportion defective of p_1 should be accepted about $100(1 - \alpha)$ percent of the time and those with proportion defective p_2 should be accepted only about 100β percent of the time.

The points $(p_1, 1 - \alpha)$ and (p_2, β) are two points through which the statistician would like to pass a curve called the operating characteristic (O.C.) curve. This curve is actually a function of the actual lot proportion defective and is defined as the probability of acceptance of a given lot with proportion defective p. It is denoted in this report as $P_a(p)$.

The difference between σ-known and σ-unknown plans is that, generally, much smaller sample sizes are needed for the σ-known plans to have them achieve essentially the same O.C. curve as the σ-unknown plan.

It is difficult to imagine a situation in which the standard deviation is ever really known. Generally what is meant by "known" is that an estimate from past data is available. In fact, it is sometimes advocated that a σ-known plan not be used unless a "good" estimate of σ is available. See, for example, page 224 of Duncan (ref. 1).

Duncan (ref. 1) considers sampling plans which meet the needs of both consumer and producer for the cases where σ is assumed to be known and also where σ is not assumed known. The results given in Duncan (ref. 1) are briefly summarized later and the extension to the case where σ is assumed unknown but constant over all lots is given.

This report then studies the procedure whereby σ-unknown plans are used for a certain number N of lots. Then a pooled estimate of σ may be made and this value used as if it were the true standard deviation. The effect of this procedure upon the operation and results of the σ-known plans is then presented. Following that, a rule for determining an acceptable value of N is given and the whole procedure is illustrated by a hypothetical example.

Special problems may arise in the double limit case if the standard deviation is large compared to the difference $X_U - X_L$. For simplification, further discussion will implicitly assume the standard deviation is not too large and all discussion will be with regard to the double limit case since it is slightly more general.

SYMBOLS

a_{α_2}	lower 100 α_2 percent point of $\chi^2_{N(n_u-1)}$ variate
$b_{1-\alpha_1}$	upper 100(1 - α_1) percent point of $\chi^2_{N(n_u-1)}$ variate
k, k^*	critical values used in acceptance test
M	number of items in a lot
MSD	maximum standard deviation
N	number of lots over which to pool estimators for $\hat{\sigma}$
$N(\mu, \sigma^2)$	normal distribution with mean μ and variance σ^2
n	number of items randomly sampled from a lot
n_k	sample size for σ-known plan
n_u	sample size for σ-unknown plan
P(E)	probability of event E
$P\left\{E_1 \mid E_2\right\}$	conditional probability of event E_1 given that event E_2 occurred
$P_a(p)$	probability of accepting a lot with true proportion defective p
$P_a^+(p_i)$	upper confidence limit on operating characteristic curve at p_i
$P_a^-(p_i)$	lower confidence limit on operating characteristic curve at p_i
$P_a(p \mid \hat{\sigma})$	conditional probability of accepting a lot with true proportion defective p, using estimate $\hat{\sigma}$

p	proportion defective
p_1	producers quality point
p_2	consumers quality point
s	sample standard deviation
v	degrees of freedom
X	specific value of random variable x
$\overline{X}$	estimate of μ
X_L	value such that if $X < X_L$, item is considered defective
X_U	value such that, if $X > X_U$, item is considered defective
x	random variable characterizing inspected item
Z_γ	$Z_\gamma = \Phi^{-1}(\gamma)$
z	denotes random variable following standard normal distribution
α	producers risk
β	consumers risk
μ	mean value or expected value of x
μ_p	value of μ such that, if $x \sim N\left(\mu_p, \sigma^2\right)$, then observed item has probability p of being defective
σ	standard deviation of x
σ^2	variance of x
$\hat{\sigma}$	estimate of σ
$\Phi(Z)$	cumulative standard normal distribution function, i.e., $P(z \leq Z) = \Phi(Z)$
$\Phi^{-1}(p)$	inverse cumulative standard normal distribution function
χ_r^2	Chi-squared distribution with r degrees of freedom
$\sim$	is distributed as
$\dot{\sim}$	is approximately distributed as

DISCUSSION OF ASSUMPTIONS

It is assumed that the output of a manufacturing process is a product whose critical value is normally distributed. It is also assumed that the variance of the product is constant while the mean varies slowly with time. An attempt should be made to justify these assumptions.

First consider the assumption that the critical variable is normally distributed. In many industrial situations the items are coming from a process that is "in control." This term means that the major sources of variability of the product have been identified and eliminated. There remains an accumulation or sum of a number of small effects. A generalization of the central limit theorem due to Lapunov gives conditions under which a sum of independent random variables approximates a normal distribution (see p. 202 of ref. 2). A very loose statement of the theorem is that a sum of m independent random variables with about the same second and third moments approaches a normal distribution as m gets large. This theorem can thus be applied to the accumulation of the small random effects to justify the normality of the critical variable.

An empirical test for the joint normality of a group of independent samples is provided by Wilk and Shapiro (ref. 3). This test does not assume equal variances of the populations nor does it assume equal means. Thus, it is an ideal test for this application. If normality is rejected, some other distribution and method must be used. To justify the assumption of equal variances, a test such as Bartlett's test for homogeneity of variance (pp. 642 to 643 of ref. 1) may be applied. If equality of variances is rejected, the procedure described here may not be used. The σ-unknown procedure should be used for all lots. It should also be recalled that the product is assumed to come from a process whose mean varies slowly with time. Thus, a procedure which uses one large initial sample to estimate σ would have a much higher probability of giving a biased estimate of σ than one using a series of independent small samples. There will be a small loss of efficiency due to the loss of one degree of freedom per sample.

σ-UNKNOWN PLANS

Suppose a double-limit variables acceptance plan is required for a situation in which σ is unknown and X_L, X_U, α, β, p_1, and p_2 are specified. Let

$$\Phi^{-1}(1 - p_1) = Z_{1-p_1}$$

and

$$\Phi^{-1}(1 - p_2) = Z_{1-p_2}$$

Then Duncan (p. 240, ref. 1) gives the following equations to obtain k^* and n_u:

$$k^* = \frac{Z_{1-\alpha}Z_{1-p_2} + Z_{1-\beta}Z_{1-p_1}}{Z_{1-\alpha} + Z_{1-\beta}} \tag{1}$$

$$n_u \cong \left[1 + \frac{(k^*)^2}{2}\right]\left(\frac{Z_{1-\alpha} + Z_{1-\beta}}{Z_{1-p_1} - Z_{1-p_2}}\right)^2 \tag{2}$$

The procedure of the test is to observe a sample of size n_u from a lot of items and compute

$$\overline{X} = \frac{\sum_{i=1}^{n_u} X_i}{n_u} \tag{3}$$

$$s = \left[\frac{\sum_{i=1}^{n_u} (X_i - \overline{X})^2}{n_u - 1}\right]^{1/2} \tag{4}$$

The lot is then accepted if

$$\left.\begin{aligned} &\frac{\overline{X} - X_L}{s} \geq k^* \\ &\frac{X_U - \overline{X}}{s} \geq k^* \\ \text{and} \quad &s \leq \text{MSD} \end{aligned}\right\} \tag{5}$$

where MSD is the maximum standard deviation as explained in Duncan (p. 246, ref. 1). If any of the previous three conditions fails to hold, the lot is rejected.

σ-KNOWN PLANS

Suppose a double-limit plan is needed when X_U, X_L, α, β, p_1, and p_2 are specified.

The parameters n_k and k are then given by

$$n_k \approx \left(\frac{Z_{1-\alpha} + Z_{1-\beta}}{Z_{1-p_1} - Z_{1-p_2}} \right)^2 \tag{6}$$

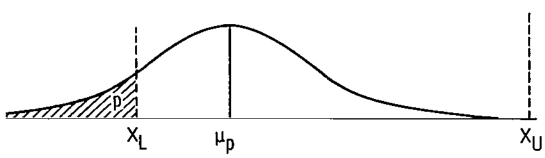
$$k = \frac{\left(Z_{1-p_1} - \frac{Z_{1-\alpha}}{\sqrt{n_k}} + Z_{1-p_2} + \frac{Z_{1-\beta}}{\sqrt{n_k}} \right)}{2} \tag{7}$$

The operation of this plan is to observe a random sample of size n_k and compute $\overline{X}$ as in equation (3). The lot is accepted if

$$\left. \begin{array}{l} \dfrac{\overline{X} - X_L}{\sigma} \geq k \\[2ex] \text{and} \\[2ex] \dfrac{X_U - \overline{X}}{\sigma} \geq k \end{array} \right\} \tag{8}$$

If either of the previous two conditions fails, the lot is rejected.

To find the O.C. curve for this plan, let p denote the true population proportion defective and temporarily assume they are at the lower limit; that is, the proportion of defectives having $X > X_U$ is negligible. This may be represented graphically as



The lot is accepted if

$$\frac{\overline{X} - X_L}{\sigma} \geq k$$

Thus,

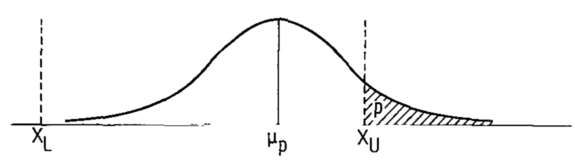
$$P_a(p) = P\left\{\frac{\overline{X} - X_L}{\sigma} \geq k\right\} = P\left\{\frac{\overline{X} - \mu_p}{\sigma} + \frac{\mu_p - X_L}{\sigma} \geq k\right\} = P\left\{\frac{\overline{X} - \mu_p}{\sigma} \geq k + \frac{X_L - \mu_p}{\sigma}\right\}$$

$$= P\left\{\frac{\overline{X} - \mu_p}{\sigma} \geq k + Z_p\right\} = P\left\{\frac{\overline{X} - \mu_p}{\sigma/\sqrt{n_k}} \geq (k + Z_p)\sqrt{n_k}\right\} \tag{9}$$

where

$$\frac{\overline{X} - \mu_p}{\sigma/\sqrt{n_k}} \sim N(0, 1)$$

If the defectives are at the upper limit, that is, the proportion of defectives having $X < X_L$ is negligible, this may be represented graphically as



and the lot is accepted if

$$\frac{X_U - \overline{X}}{\sigma} \geq k$$

Then

$$P_a(p) = P\left\{\frac{X_U - \overline{X}}{\sigma} \geq k\right\} = P\left\{\frac{X_U - \mu_p}{\sigma} - \frac{\overline{X} - \mu_p}{\sigma} \geq k\right\} = P\left\{\frac{\overline{X} - \mu_p}{\sigma} \leq \frac{X_U - \mu_p}{\sigma} - k\right\}$$

$$= P\left\{\frac{\overline{X} - \mu_p}{\sigma} \leq Z_{1-p} - k\right\} = P\left\{\frac{\overline{X} - \mu_p}{\sigma} \leq - Z_p - k\right\} = P\left\{\frac{\overline{X} - \mu_p}{\sigma/\sqrt{n_k}} \leq (-k - Z_p)\sqrt{n_k}\right\}$$

which by symmetry of the standard normal distribution becomes

$$P_a(p) = P\left\{\frac{\overline{X} - \mu_p}{\sigma/\sqrt{n_k}} \geq (k + Z_p)\sqrt{n_k}\right\}$$

since

$$\frac{\overline{X} - \mu_p}{\sigma/\sqrt{n_k}} \sim N(0, 1)$$

σ-ESTIMATED PLANS

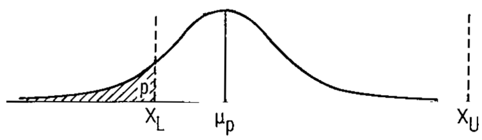
Noting equations (6) and (7), it is readily seen that knowledge of σ is not necessary to specify the parameters of the plan. Denote the estimate of σ by $\hat{\sigma}$. Then in place of equation (8) we get

$$\left.\begin{array}{l} \dfrac{\overline{X} - X_L}{\hat{\sigma}} \geq k \\ \\ \text{and} \\ \\ \dfrac{X_U - \overline{X}}{\hat{\sigma}} \geq k \end{array}\right\} \qquad (10)$$

Let the particular estimate $\hat{\sigma}$ be given by

$$\hat{\sigma} = \left[\frac{\sum_{j=1}^{N} \sum_{i=1}^{n_u} \left(X_{ij} - \overline{X}_j \right)^2}{N(n_u - 1)} \right]^{1/2} \tag{11}$$

To find the O. C. curve for this plan, let p denote the true population proportion defective and assume the defectives are at the low end. This is represented graphically as



The lot is accepted if

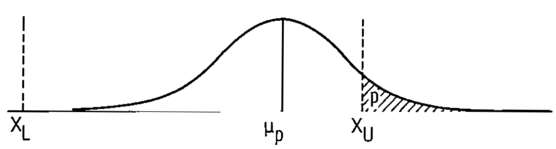
$$\frac{\overline{X} - X_L}{\hat{\sigma}} \geq k$$

The probability of this event depends upon the value of $\hat{\sigma}$ and is given by

$$P_a(p|\hat{\sigma}) = P\left\{ \frac{\overline{X} - X_L}{\hat{\sigma}} \geq k \middle| \hat{\sigma} \right\} = P\left\{ \frac{\overline{X} - X_L}{\sigma} \geq k \frac{\hat{\sigma}}{\sigma} \middle| \hat{\sigma} \right\} = P\left\{ \frac{\overline{X} - \mu_p}{\sigma} - \frac{X_L - \mu_p}{\sigma} \geq k \frac{\hat{\sigma}}{\sigma} \middle| \hat{\sigma} \right\}$$

$$= P\left\{ \frac{\overline{X} - \mu_p}{\sigma} \geq k \frac{\hat{\sigma}}{\sigma} + Z_p \middle| \hat{\sigma} \right\} = P\left\{ \frac{\overline{X} - \mu_p}{\sigma / \sqrt{n_k}} \geq \left(k \frac{\hat{\sigma}}{\sigma} + Z_p \right) \sqrt{n_k} \middle| \hat{\sigma} \right\} \tag{12}$$

If the defectives are at the upper end, we have the following graphic representation:



$$P_a(p|\hat{\sigma}) = P\left\{\frac{X_U - \overline{X}}{\hat{\sigma}} \geq k \middle| \hat{\sigma}\right\} = P\left\{\frac{X_U - \mu_p}{\sigma} - \frac{\overline{X} - \mu_p}{\sigma} \geq k \frac{\hat{\sigma}}{\sigma} \middle| \hat{\sigma}\right\}$$

$$= P\left\{\frac{\overline{X} - \mu_p}{\sigma} \leq \frac{X_U - \mu_p}{\sigma} - k \frac{\hat{\sigma}}{\sigma} \middle| \hat{\sigma}\right\} = P\left\{\frac{\overline{X} - \mu_p}{\sigma} \leq -Z_p - k \frac{\hat{\sigma}}{\sigma} \middle| \hat{\sigma}\right\}$$

$$= P\left\{\frac{\overline{X} - \mu_p}{\sigma/\sqrt{n_k}} \geq \left(Z_p + k \frac{\hat{\sigma}}{\sigma}\right) \sqrt{n_k} \middle| \hat{\sigma}\right\}$$

and the result is identical to equation (12).

It is now seen that the conditional probability of acceptance of lots, after the estimate $\hat{\sigma}$ is made and used, is a random variable whose distribution depends upon the distribution of $\hat{\sigma}$. That is, the actual probability of acceptance is unknown but constant for all lots following the first N lots, and for different initial samples of N lots it would vary.

DETERMINATION OF CONFIDENCE LIMITS AND N

It is well known that

$$\frac{N(n_u - 1)\hat{\sigma}^2}{\sigma^2} \sim \chi^2_{N(n_u-1)}$$

Thus, let

$$P\left\{a_{\alpha_2} \leq \frac{N(n_u - 1)\hat{\sigma}^2}{\sigma^2} \leq b_{1-\alpha_1}\right\} = P\left\{\sqrt{\frac{a_{\alpha_2}}{N(n_u - 1)}} \leq \frac{\hat{\sigma}}{\sigma} \leq \sqrt{\frac{b_{1-\alpha_1}}{N(n_u - 1)}}\right\} = 1 - \alpha_1 - \alpha_2$$

where both square rocts are assumed to be the positive values and where a_{α_2} is the lower $100 \times \alpha_2$ percent and $b_{1-\alpha_1}$ the upper $100 \times (1 - \alpha_1)$ percent points of a $\chi^2_{N(n_u-1)}$ variate.

Then, given N, α_1, and α_2, upper and lower limits for the O.C. curve could be calculated at any point p from equation (12) and a set of tables of the chi-square and normal distribution functions.

Now the same upper and lower confidence limits are not necessarily required for all p. Denote upper and lower confidence limits on the O.C. curve at $p = p_i$ where $i = 1, 2$ by $P_a^+(p_i)$ and $P_a^-(p_i)$, respectively.

The same confidence levels are not necessarily required for all values of p either. In fact, a producer has no concern over an overestimated O.C. curve at p_1 and a consumer has no concern over an underestimated O.C. curve at p_2.

Thus, let $\alpha_{1,p}$ denote the lower confidence level desired on the lower limit of the O.C. curve at the point p and $\alpha_{2,p}$ denote the upper confidence level desired on the upper confidence limit of the O.C. curve at the point p.

From a set of tables of the chi-square distribution we could determine by trial and error a minimum N that would satisfy

$$P\left\{P_a^-(p_i) \leq P_a\left(p_i \middle| \hat{\sigma}\right) \leq P_a^+(p_i)\right\} \geq 1 - \left(\alpha_{1, p_i} + \alpha_{2, p_i}\right) \qquad \text{for } i = 1, 2 \tag{13}$$

If N satisfied the inequalities exactly, the situation could be represented as in figure 2 where the dash dot curve represents the nominal O.C. curve and the dashed upper and lower curves denote upper and lower confidence limits. Density functions of the attained O.C. curve are represented vertically on the dashed lines at $p = p_1$ and $p = p_2$.

The determination of N may be simplified if we require that $N(n_u - 1) > 30$. This would most likely be true in practical situations. Then we may use the fact (ref. 4, p. 400) that

$$\sqrt{2\chi_v^2} \doteq N\left(\sqrt{2v - 1},\ 1\right)$$

for $v > 30$. But we have

$$\frac{N(n_u - 1)\hat{\sigma}^2}{\sigma^2} \sim \chi^2_{N(n_u - 1)}$$

so that

$$\frac{\hat{\sigma}}{\sigma}\sqrt{2N(n_u - 1)} \doteq N\left(\sqrt{2N(n_u - 1) - 1},\ 1\right)$$

From equation (13) we get

$$P\left\{P_a^-(p_i) \le P_a\left(p_i \mid \hat{\sigma}\right) \le P_a^+(p_i)\right\} = P\left\{P_a^-(p_i) \le P\left\{z \ge \left(k\frac{\hat{\sigma}}{\sigma} + Z_{p_i}\right)\sqrt{n_k} \,\middle|\, \hat{\sigma}\right\} \le P_a^+(p_i)\right\}$$

$$= P\left\{P_a^-(p_i) \le P\left\{z \le \left(-Z_{p_i} - k\frac{\hat{\sigma}}{\sigma}\right)\sqrt{n_k} \,\middle|\, \hat{\sigma}\right\} \le P_a^+(p_i)\right\}$$

$$= P\left\{Z_i^- \le \left(Z_{1-p_i} - k\frac{\hat{\sigma}}{\sigma}\right)\sqrt{n_k} \le Z_i^+\right\} \tag{14}$$

where z is a unit normal variate and where

$$Z_i^- = \Phi^{-1}\left(P_a^-(p_i)\right)$$

$$Z_i^+ = \Phi^{-1}\left(P_a^+(p_i)\right)$$

Let

$$z = \left(Z_{1-p_i} - k\frac{\hat{\sigma}}{\sigma}\right)\sqrt{n_k}$$

Recalling that $\omega \sim N(\mu, \sigma^2)$ implies $a + b\omega \sim N(a + b\mu, b^2\sigma^2)$, it is easily seen that

$$z = Z_{1-p_i}\sqrt{n_k} - k\frac{\hat{\sigma}}{\sigma}\sqrt{n_k} = Z_{1-p_i}\sqrt{n_k} + \left[\frac{-k\sqrt{n_k}}{\sqrt{2N(n_u - 1)}}\right]\left[\frac{\hat{\sigma}}{\sigma}\sqrt{2N(n_u - 1)}\right]$$

$$= a_i + b\left[\frac{\hat{\sigma}}{\sigma}\sqrt{2N(n_u - 1)}\right]$$

where

$$\left[\frac{\hat{\sigma}}{\sigma}\sqrt{2N(n_u - 1)}\right] \mathrel{\dot{\sim}} N\left(\sqrt{2N(n_u - 1) - 1},\ 1\right)$$

implies $z \mathrel{\dot{\sim}} N\left(a_i + b\sqrt{2N(n_u - 1) - 1},\ b^2\right)$. Hence,

$$z^* = \frac{z - a_i - b\sqrt{2N(n_u - 1) - 1}}{|b|} \sim N(0, 1)$$

and equation (14) may be rewritten as

$$P\left\{Z_i^- \le z \le Z_i^+\right\} = P\left\{\frac{Z_i^- - a_i - b\sqrt{2N(n_u - 1) - 1}}{|b|} \le z^* \le \frac{Z_i^+ - a_i - b\sqrt{2N(n_u - 1) - 1}}{|b|}\right\} \tag{15}$$

The problem is thus solved if an N can be found to satisfy all of the following simultaneously:

$$Z_{\alpha_{1}, p_1} \ge \frac{Z_1^- - a_1 - b\sqrt{2N(n_u - 1) - 1}}{|b|} \tag{16.1}$$

$$Z_{1-\alpha_{2}, p_1} \le \frac{Z_1^+ - a_1 - b\sqrt{2N(n_u - 1) - 1}}{|b|} \tag{16.2}$$

$$Z_{\alpha_{1}, p_2} \ge \frac{Z_2^- - a_2 - b\sqrt{2N(n_u - 1) - 1}}{|b|} \tag{16.3}$$

$$Z_{1-\alpha_{2}, p_2} \le \frac{Z_2^+ - a_2 - b\sqrt{2N(n_u - 1) - 1}}{|b|} \tag{16.4}$$

To derive an approximating solution, let N_i be the minimum N which will satisfy equation (16.i) where i = 1, 2, 3, or 4. These inequalities will then become approximate equalities. Solving these for N_i and then setting $N = \max\{N_i\}$ will yield the desired solution. For example, let i = 1. Then we have

$$Z_{\alpha_{1}, p_1} |b| \approx Z_1^- - a_1 - b\sqrt{2N(n_u - 1) - 1}$$

Thus,

$$\frac{Z_{\alpha_{1},p_1} k\sqrt{n_k}}{\sqrt{2N_1(n_u - 1)}} = Z_1^- - Z_{1-p_1}\sqrt{n_k} + \frac{k\sqrt{n_k}\left[2N_1(n_u - 1) - 1\right]}{\sqrt{2N_1(n_u - 1)}} \tag{17}$$

For large values of $N_1(n_u - 1)$, it is seen that

$$2N_1(n_u - 1) \approx 2N_1(n_u - 1) - 1$$

Then let

$$N_1^* = 2N_1(n_u - 1) \approx 2N_1(n_u - 1) - 1$$

and substitute this in equation (17) to get

$$\frac{Z_{\alpha_{1},p_1} k\sqrt{n_k}}{\sqrt{N_1^*}} \approx Z_1^- - Z_{1-p_1}\sqrt{n_k} + k\sqrt{n_k} \tag{18}$$

or

$$N_1^* \approx \left(\frac{Z_{\alpha_{1},p_1} k\sqrt{n_k}}{Z_1^- - Z_{1-p_1}\sqrt{n_k} + k\sqrt{n_k}}\right)^2$$

or

$$N_1 \approx \frac{1}{2(n_u - 1)}\left[\frac{Z_{\alpha_{1},p_1} k\sqrt{n_k}}{Z_1^- - \left(Z_{1-p_1} - k\right)\sqrt{n_k}}\right]^2 \tag{19}$$

Similar derivations beginning with equations (16.2), 16.3), and (16.4) will yield

$$N_2 \approx \frac{1}{2(n_u - 1)} \left[\frac{Z_{1-\alpha_2, p_1} k\sqrt{n_k}}{Z_1^+ - \left(Z_{1-p_1} - k\right)\sqrt{n_k}} \right]^2 \tag{20}$$

$$N_3 \approx \frac{1}{2(n_u - 1)} \left[\frac{Z_{\alpha_1, p_2} k\sqrt{n_k}}{Z_2^- - \left(Z_{1-p_2} - k\right)\sqrt{n_k}} \right]^2 \tag{21}$$

$$N_4 \approx \frac{1}{2(n_u - 1)} \left[\frac{Z_{1-\alpha_2, p_2} k\sqrt{n_k}}{Z_2^+ - \left(Z_{1-p_2} - k\right)\sqrt{n_k}} \right]^2 \tag{22}$$

EXAMPLE

Suppose a consumer purchases large quantities of rivets. If a rivet is larger in diameter than 0.252 centimeter, it will not fit in the standard size rivet hole and if it is smaller than 0.240 centimeter it will slip through the standard size rivet hole.

The producer of the rivets feels very strongly that his process is in very tight control and thus he can achieve the limitations of $X_U = 0.252$ centimeter and $X_L = 0.240$ centimeter very easily. The consumer is not ready to accept this on faith so he will institute a variables acceptance sampling plan. Both agree that the diameter of rivets is normally distributed.

The consumer feels that if there is much over 15 percent defective items per lot, his workers will be wasting too much time with defects for him to readily absorb. He also feels that a level of 1 percent defective is desired. Thus the consumer and producer mutually agree upon the following nominal figures:

$$p_1 = 0.01 \qquad \alpha = 0.05$$

$$p_2 = 0.15 \qquad \beta = 0.10$$

Then

$$Z_{1-p_1} = \Phi^{-1}(1 - p_1) = 2.326 \qquad Z_{1-\alpha} = 1.645$$

$$Z_{1-p_2} = \Phi^{-1}(1 - p_2) = 1.036 \qquad Z_{1-\beta} = 1.282$$

and

$$k^* = \frac{Z_{1-\alpha}Z_{1-p_2} + Z_{1-\beta}Z_{1-p_1}}{Z_{1-\alpha} + Z_{1-\beta}} = \frac{(1.645)(1.036) + (1.282)(2.326)}{1.645 + 1.282} = 1.601$$

$$n_u \approx \left[1 + \frac{(k^*)^2}{2}\right]\left(\frac{Z_{1-\alpha} + Z_{1-\beta}}{Z_{1-p_1} - Z_{1-p_2}}\right)^2 = \left[1 + \frac{(1.601)^2}{2}\right]\left(\frac{1.645 + 1.282}{2.326 - 1.036}\right)^2 = 11.75 \cong 12$$

The producer and consumer also mutually agree that the actual probability of acceptance at p_1 should be no less than 0.988 with probability 0.999 and the actual probability of acceptance at p_2 should be no greater than 0.11 with probability 0.999. Thus,

$$P_a^-(p_1) = 0.988 \qquad \alpha_{1,p_1} = 0.001$$

$$P_a^+(p_2) = 0.11 \qquad \alpha_{2,p_2} = 0.001$$

Only equations (19) and (22) really need be considered since an overestimate of the O.C. curve at p_1 and an underestimate at p_2 are not really undesirable consequences. The following quantities are needed for the calculations:

$$Z_{\alpha_{1,p_1}} = Z_{0.001} = -3.090 \qquad Z_{p_1} = Z_{0.01} = -2.326$$

$$Z_{1-\alpha_{2,p_2}} = Z_{0.999} = 3.090 \qquad Z_{p_2} = Z_{0.15} = -1.036$$

$$Z_1^- = \Phi^{-1}(0.988) = 2.257 \qquad Z_2^+ = \Phi^{-1}(0.11) = -1.226$$

Thus, from equations (6) and (7)

$$n_k = \left(\frac{1.645 + 1.282}{2.326 - 1.036}\right) = 2.26 \approx 3$$

$$k = \frac{\left(2.326 - \frac{1.645}{\sqrt{3}}\right) + \left(1.036 + \frac{1.282}{\sqrt{3}}\right)}{2} = 1.576$$

Then

$$N_1 = \frac{1}{2(11)}\left[\frac{(-3.090)(1.576)\sqrt{3}}{2.257 - (2.326 - 1.576)\sqrt{3}}\right]^2 = 3.52 \approx 4 \qquad (23)$$

$$N_4 = \frac{1}{2(11)}\left[\frac{(3.090)(1.576)\sqrt{3}}{-1.226 - (1.036 - 1.576)\sqrt{3}}\right]^2 = 38.27 \approx 38$$

Thus a conservative procedure is to pool information over 38 lots, estimate σ, and then use the σ-known plan.

CONCLUSION

A method has been presented for evaluating the effect upon the design and operating characteristic curve of a standard deviation σ known acceptance sampling plan when in fact the standard deviation has been estimated from previous data.

It was shown that the design of the standard deviation known plan is unaffected, but that the operating characteristic (O. C.) curve becomes a random variable. It was then shown how to use an approximation based upon the normal distribution to compute the number of observations necessary to estimate σ such that the actual O. C. curve falls within specified confidence limits of the nominal O. C. curve.

Lewis Research Center,
National Aeronautics and Space Administration,
Cleveland, Ohio, January 27, 1971,
129-04.

REFERENCES

1. Duncan, Acheson J.: Quality Control and Industrial Statistics. Third ed., R. D. Irwin, Inc., 1965.

2. Fisz, Marek: Probability Theory and Mathematical Statistics. Third ed., John Wiley & Sons, Inc., 1963.

3. Wilk, Martin B.; and Shapiro, S. S.: The Joint Assessment of Normality of Several Independent Samples. Technometrics, vol. 10, no. 4, Nov. 1968, pp. 825-840.

4. Kendall, Maurice G.; and Stuart, Alan: The Advanced Theory of Statistics. Vol. 1. Second ed., Griffin & Co., Ltd., 1963.

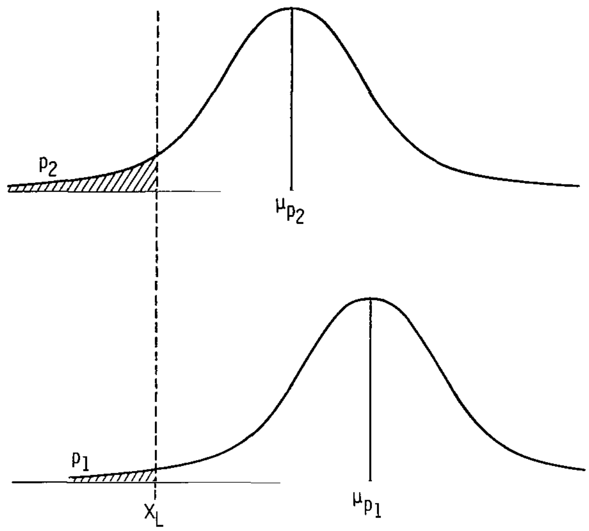


Figure 1. - Normal density functions with proportion defective p_2 (top) and p_1 (bottom).

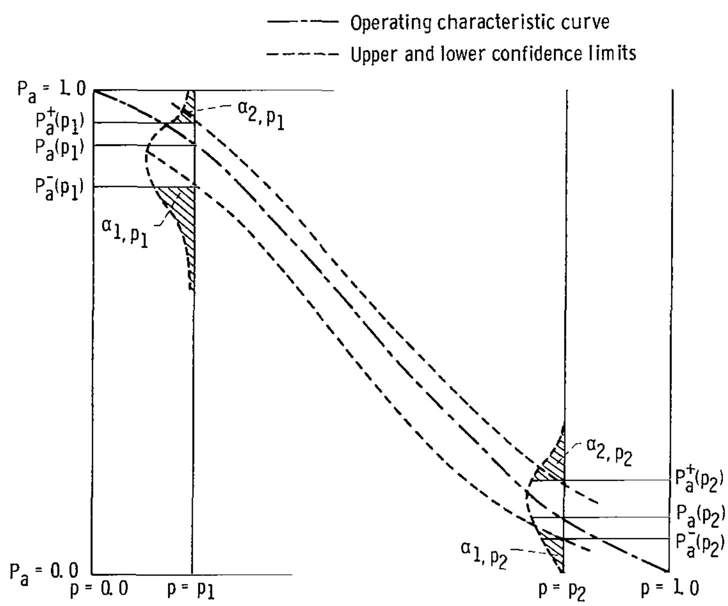


Figure 2. - Possible nominal O.C. curve with lower and upper confidence limits. (Possible density functions of attained O.C. curve at p_1 and p_2 are sketched vertically.)

NASA-Langley, 1971 —— 19 E-6034

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
WASHINGTON, D.C. 20546

OFFICIAL BUSINESS
PENALTY FOR PRIVATE USE $300

FIRST CLASS MAIL

POSTAGE AND FEES PAID
NATIONAL AERONAUTICS AND
SPACE ADMINISTRATION

02U 001 44 51 3DS 71088 00903
AIR FORCE WEAPONS LABORATORY /WLOL/
KIRTLAND AFB, NEW MEXICO 87117

ATT E. LOU BOWMAN, CHIEF,TECH. LIBRARY

POSTMASTER: If Undeliverable (Section 158 Postal Manual) Do Not Return

"The aeronautical and space activities of the United States shall be conducted so as to contribute . . . to the expansion of human knowledge of phenomena in the atmosphere and space. The Administration shall provide for the widest practicable and appropriate dissemination of information concerning its activities and the results thereof."

—NATIONAL AERONAUTICS AND SPACE ACT OF 1958

NASA SCIENTIFIC AND TECHNICAL PUBLICATIONS

TECHNICAL REPORTS: Scientific and technical information considered important, complete, and a lasting contribution to existing knowledge.

TECHNICAL NOTES: Information less broad in scope but nevertheless of importance as a contribution to existing knowledge.

TECHNICAL MEMORANDUMS: Information receiving limited distribution because of preliminary data, security classification, or other reasons.

CONTRACTOR REPORTS: Scientific and technical information generated under a NASA contract or grant and considered an important contribution to existing knowledge.

TECHNICAL TRANSLATIONS: Information published in a foreign language considered to merit NASA distribution in English.

SPECIAL PUBLICATIONS: Information derived from or of value to NASA activities. Publications include conference proceedings, monographs, data compilations, handbooks, sourcebooks, and special bibliographies.

TECHNOLOGY UTILIZATION PUBLICATIONS: Information on technology used by NASA that may be of particular interest in commercial and other non-aerospace applications. Publications include Tech Briefs, Technology Utilization Reports and Technology Surveys.

Details on the availability of these publications may be obtained from:

SCIENTIFIC AND TECHNICAL INFORMATION OFFICE

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

Washington, D.C. 20546